MASSACHUSETTS INSTITUTE of TECHNOLOGY
LIBRARIES

Digitized by the Internet Archive
in 2011 with funding from
Boston Library Consortium Member Libraries

http://www.archive.org/details/mentoringdiversi00athe

HB31
.M415
no.98-2

DEWEY

working paper
department
of economics

MENTORING AND DIVERSITY

Susan Athey
Christopher Avery
Peter Zensky

No. 98-2 January, 1998

massachusetts
institute of
technology
50 memorial drive
cambridge, mass. 02139

WORKING PAPER
DEPARTMENT
OF ECONOMICS

MENTORING AND DIVERSITY

Susan Athey
Christopher Avery
Peter Zensky

No. 98-2 January, 1998

MASSACHUSETTS
INSTITUTE OF
TECHNOLOGY

50 MEMORIAL DRIVE
CAMBRIDGE, MASS. 02142

MASSACHUSETTS INSTITUTE
OF TECHNOLOGY
JAN 2 8 1999
LIBRARIES

Mentoring and Diversity

Susan Athey Christopher Avery Peter Zemsky*

December, 1997

Abstract

This paper studies the forces which determine how diversity at a firm evolves over time. We consider a dynamic model of a single firm with two levels of employees, the entry level and the upper level. In each period, the firm selects a subset of the entry-level workers for promotion to the upper level. The members of the entry-level worker pool vary in their initial ability as well as in their "type," where type could refer to gender or cultural background. Employees augment their initial ability by acquiring specific human capital in mentoring interactions with upper level employees. We assume that an entry-level worker receives more mentoring when a greater proportion of upper-level workers match the entry-level worker's type. In this model, it is optimal for the firm to consider type in addition to ability in making promotion decisions, so as to maximize the effectiveness of future mentoring. We derive conditions under which firms attain full diversity, as well as conditions under which there are multiple steady states, so that the level of diversity depends on the firm's initial conditions. With multiple steady states, temporary affirmative action policies can have a long-run impact on diversity levels.

*MIT, Kennedy School of Government, and INSEAD. We would like to thank Preston McAfee and two anonymous referees for insightful comments which greatly improved the paper. We are also grateful to Jim Baron, Jeremy Bulow, Peter Diamond, Joshua Gans, Ed Lazear, Michael Kremer, Bill Lovejoy, Meg Meyer, Bob Marshall, John Roberts, Scott Stern, Bob Wilson, and seminar participants at Stanford and Oxford for helpful discussions. All errors remain our own. The authors acknowledge support from the State Farm Companies Foundation and the National Science Foundation (Athey, SBR-9631760). Athey and Avery are also grateful to the Cowles Foundation at Yale for hospitality and support while this work was being completed.

1 Introduction

The topics of workplace diversity and of biases in hiring and promotion decisions engender impassioned discourse in contemporary American society. Some complain that affirmative action policies lead firms to discriminate against white males. Others counter that the historic domination of management by white males puts other groups at a disadvantage, arguing that the underlying nature of the workplace inherently favors those who are from similar backgrounds as their managers. They point to the fact that in many industries and occupations, women and minorities have moved only slowly into high-level positions, despite the fact that the labor pool and lower levels of the workforce have been diverse for some time.[1] This phenomenon has been referred to as the "glass ceiling."

Some firms have made active efforts to reshape themselves, instituting affirmative action programs and hiring diversity managers. For example, IBM is known for its affirmative action program, while American Airlines, AT&T, Colgate-Palmolive Corp., DuPont Corp., and Pacific Bell have instituted diversity programs.[2] Other firms aggressively fight external constraints on their staffing decisions. It is difficult to reconcile this heterogeneity among firms without specifying why firms care about diversity.

This paper develops a model to analyze the relationship between the diversity of a firm's upper level employees and its internal promotion policies. We consider the dynamic problem faced by a firm in choosing which of its lower level employees to promote as its existing upper-level workers retire.[3] Employees are characterized by an initial ability for upper level work and by a type, which can be interpreted as gender, race, cultural background, or personality type. Employees augment their initial ability by acquiring specific human capital in mentoring interactions with upper level employees. We interpret mentoring broadly, including activities such as information sharing, informal teaching, or career advice provided by more senior workers.

The critical assumption in our model is that an entry level employee acquires more human capital from mentoring when the firm has more upper-level employees who match her type. This assumption is consistent with evidence from a variety of sources. Psychologists and sociologists have documented that mentoring relationships within

[1] In many occupations, the 1980 Census revealed a higher proportion of women, Blacks, and Hispanics in "employee" positions than in "supervisor" positions (Rothstein, 1997). Large gaps exist between the proportion of women at lower levels and higher levels of large corporations (Morrison and VanGlinow, 1990) and law firms (Spurr, 1990), and also to a lesser extent in academic economics departments (Bartlett, 1997).

[2] *The Wall Street Journal* (August 9, 1991, p. B1).

[3] Our model can also be used to study the dynamically optimal *hiring* policy of a firm where the productivity of new hires depends in part on the proportion of more senior workers matching their type. Under this interpretation, the analysis addresses the evolution of the diversity of a firm's entire workforce. For clarity, we limit our exposition to the promotions interpretation, except for a discussion of hiring in Section 2.1.1.

firms are more likely to form between members of the same group.[4] Ibarra (1992) demonstrates that the structure of social networks depends on gender and race. More generally, communication, and thus mentoring, may be more natural and more effective when people share common interests (such as sports), cultural experiences, language, or when people have significant interactions in a community outside the workplace.[5] Type-biased mentoring has also been highlighted in accounts of business and the legal profession. For example, a *National Law Journal* article[6] reported, "Despite women's progress, the partnership ranks are 86.4% male. This puts women at a disadvantage when it comes to mentoring, the most important factor in becoming a partner. Often it's harder for women to find a mentor, because the massive majority of partners are men, and men tend to be more comfortable mentoring other men. Rainmaking and client development—skills typically learned from a mentor—are keys to partnership."[7]

Our analysis explores the consequences of type-based mentoring for promotion policies and for the long-run evolution of diversity in a given firm. In our model, there are two types of employees, and that one type is currently the "majority" (i.e. more upper level employees are of that type). Our assumptions about mentoring imply that entry-level employees of the majority type acquire more human capital, and thus firms who base promotions solely on ability promote more majority employees. However, since the diversity of management affects a firm's profits through the human capital acquisition of future workers, the optimal policy of a far-sighted firm will generally involve promoting workers who do not have the highest ability in order to influence the evolution of diversity over time. Thus, our approach suggests that observed differences in promotion decisions result from (i) true productivity differences, which arise (partly) as a result of a firm's past promotion decisions, and (ii) what we call *bias* in

[4]Morrison and Van Glinow (1990) and Noe (1988) review the theory and evidence in favor of type-biased mentoring. More recently, Dreher and Cox (1996) document such differences in a survey of MBA graduates. Kanter's classic (1977) analysis of gender roles in organizations takes a similar view, observing that "A homogeneous network reinforced the inability of its members to incorporate heterogeneous elements" (p. 59), and further, "...numbers—proportional representation are important not only because they symbolize the presence or absence of discrimination but also because they have real consequences for performance" (p. 6).

[5]In a related set of historical examples, Greif (1993, 1994) shows that cooperation in trading relationships can at times be more easily sustained between members of an extended family or community.

[6]"Women's Progress Slows at Top Firms," May 6, 1996, p.A1.

[7]Similar claims have been made about the management consulting industry: "The lack of role models and mentors who are black is a problem...the same issues that face blacks in corporate America—lack of connections, lack of mentors, and the often-present glass ceiling—face blacks in management consulting" (Prakash, Gautam, *The Insider's Guide to Management Consulting*, San Francisco: Wet Feet Press, 1995). Similarly, *The Wall Street Journal* ("Women Make Strides, But Men Stay Firmly in Top Company Jobs," March 29, 1994, p. A1), reporting on the effects of culture in large U.S. companies, argued that "as long as the percentages of male managers remain high, the culture remains mostly male and, women say, indifferent or hostile to their advancement. Women say they are ignored, not taken seriously...[so] there is little chance of breaking through the glass ceiling."

promotions, that is, a decision to pass over some workers with higher ability as part of a long-term plan to move to a desired level of diversity. While much of the political debate about affirmative action and discrimination centers on issues of fairness in the evaluation of individual workers at a given point in time, our model suggests that firms can benefit from considering today's promotion decisions as an investment in future diversity.

In this context, we address the following question: Given an initial level of diversity, under what conditions do far-sighted firms benefit from increasing or decreasing diversity? In our model, the opportunity cost of a homogeneous upper level is an inability to take advantage of the scarce talent of entry-level minority employees. Since minorities receive relatively little mentoring, even a minority with a very high initial ability may be passed over in favor of a less initially able, but better mentored, majority type. We show that the more rapidly the initial ability of each type of worker diminishes in the number of that type promoted, the greater the cost of homogeneity and the more the firm will shift its bias towards the current minority type. Consistent with this analysis, the desire to exploit the talents of an increasingly diverse labor market is often cited by firms that actively promote workforce diversity.[8] Scarcity is likely to be a factor when a job requires specialized skills, for example, academics, athletics, high-tech managers, or specialist physicians.

In general, we find that the optimal bias need not favor the minority, even if there are decreasing returns to having more mentors of a given type. Because majority employees are better mentored, their promotion rates can be higher than those of minorities. As a result, the firm cares more about the effective mentoring of majority than minority employees. A profit-maximizing firm will only bias its promotions to favor increased diversity if there are sufficiently decreasing returns to mentors of a given type.

We further consider the factors which affect the long-run steady state level of diversity at a firm. We characterize conditions under which there is a unique long-run level of diversity, which may involve either full diversity or complete homogeneity. When there is a unique long-run outcome, history determines which group is in the majority but not the level of diversity. However, we also characterize conditions under which there may be multiple steady state levels of diversity. In particular, we show that even if the per-period profits of given firm may be highest if it is in the state of full diversity, the firm might not choose to move towards diversity if its starting point is

[8]For example, *Business Week* ("White, Male, and Worried," January 31, 1994, pp. 50-55) recently reported that to some companies, "managing diversity...is a competitive weapon that helps the company capitalize on its talent pool. The goal: to create a culture that enables all employees to contribute their full potential to the company's success. One way is to groom more qualified women and minorities through active succession planning." Similarly, it has been claimed that companies use diversity programs to "attract and retain the best and the brightest...giving a headstart in recruiting and managing the workers of tomorrow." (*USA Today,* "Setting diversity's foundation in the bottom line," by Del Jones, October 15, 1996, p. 4B.)

homogeneous. When there are multiple steady states, history matters in the strongest sense: the past determines both the level of diversity and which type is in the majority. Then, short-run pressure on the firm to diversify can have long lasting impact on the level of diversity by moving the firm from one steady state to another.

Our paper differs from previous theoretical work on discrimination (e.g. Arrow, 1973; Coate and Loury, 1992; Cornell and Welch, 1996; Rosen, 1997) in several ways. First, prior work is best suited for understanding discrimination in hiring rather than in promotion decisions. For example, the assumptions of incomplete or asymmetric information about worker abilities that underlies much prior work are less palatable when an employee has been with a firm for a long period of time. Second, prior work does not consider dynamically optimal policies. Finally, most studies focus on market-wide causes and remedies for the observed inequities; only a few model the forces which lead to discrimination within a firm, and all prior work deal with a one-on-one relationship between the worker and the firm. None formalizes the idea that the current diversity of a firm affects the career paths of new employees.

The policy implications from our model differ from the existing literature as well. Past work on affirmative action policies have suggested that they can actually impede the progress of minorities by reinforcing beliefs that minorities are less qualified (Coate and Loury (1992), Cornell and Welch (1996)). Here we show that affirmative action may increase diversity in the long run by inducing the firm to shift from one steady state level of diversity to a different, more diverse, steady state. Once the firm has sufficiently adjusted its level of diversity, an affirmative action policy is no longer necessary; the firm will choose to maintain full diversity of its own accord. This result derives from the fact that the short-term costs of moving towards diversity are highest for initially homogeneous firms, in which minorities face the largest mentoring disadvantage. When the initial level of diversity (and thus the gain to workers from mentoring) is more balanced, firms are more willing to move towards full diversity.

In order to fully characterize the dynamically optimal policy of a given firm, we have abstracted from a variety of factors, most notably the issue of market equilibrium. An important question is whether type-based mentoring necessarily leads to completely homogeneous firms when multiple firms compete for the same pool of entry-level workers. Of course, if the number of firms is limited and the talents of workers are specific to the firm-worker match, or if there are labor market frictions such as search costs, firms will face diverse labor pools in equilibrium, and the forces analyzed in our basic model will still play an important role for firms. Furthermore, Section 5 considers a static, multi-firm extension of our model that incorporates a labor market for entry level employees to show that under our assumptions, each firm faces diminishing returns to majority-type workers in the entry level, since only a subset of entry-level workers will be promoted. Thus, even when labor market frictions are not present, firms will not necessarily specialize by type in their entry-level workforce.

While the focus of this paper is on a firm's choice about ethnic or gender diversity

in its workforce, the model we develop could also apply to other types of diversity, such as diversity in skills (operations versus marketing skills for managers, theorists versus empiricists) or personality (cooperative versus competitive). Further, our results can also be applied to a class of production decisions about practices (such as research and development) whose productivity depends on their extent of use in the recent past.

The paper proceeds as follows. We introduce and discuss the formal model in Section 2. Section 3 characterizes when the optimal bias favors diversity or homogeneity. In Section 4 we analyze how diversity evolves over time. Section 5 extends the model to include a labor market for entry level employees. Section 6 discusses the interaction between type-based mentoring and ex ante human capital acquisition. Section 7 concludes.

2 The Model

This section introduces a simple model of a single firm managing how the diversity of its upper level employees evolves over time. The firm employs a continuum of upper level managers and a continuum of lower level employees.[9] We normalize the measure of managers at the firm to 1. There are two types of employees, labeled A and B. The proportion of managers of type A in period t is denoted m^t. In each period a proportion $r \in (0, 1]$ of the managers retire and must be replaced. Retirement rates are uniform, meaning that rm^t workers of type A and $r(1 - m^t)$ workers of type B retire in period t.

The firm operates an internal labor market in which retiring managers are replaced by lower level employees. The lower level is assumed to be integrated with an equal number of type A's and B's.[10] We assume that there are at least r entry-level workers of each type and that the two pools of workers are symmetric in their initial abilities. Member $\theta \in [0, r]$ of a pool of entry-level workers has an initial ability $x(\theta)$, which represents the surplus received by the firm when worker θ takes an upper-level position. There are several ways to justify the presence of such a surplus. The worker's ability might be specific to the firm-worker match, or it might not be publicly observable after a worker is attached to a firm. Frictions in the labor market such as search costs might also lead to such a surplus.

The function $x(\theta)$ is assumed to be nonincreasing with a continuous first derivative. Thus, the quality of the marginal worker from each pool (weakly) decreases as the firm digs deeper into that pool, representing scarcity of managerial ability among entry-level

[9] In our working paper we consider a model of type-based mentoring with a finite number of employees where initial ability is stochastic. The qualitative insights of the stochastic model are the same, but the deterministic model analyzed in this paper is more tractable.

[10] Section 5 considers an extension where the diversity of the entry level is endogenously determined through labor market interactions between firms.

workers. Such scarcity may be especially relevant for jobs in the service sector, jobs requiring education and specialized skills, jobs where "stars" are important (such as academics, entertainment, and sports), and high level management positions in firms.

A worker's type and initial ability are observable. Each applicant gains additional management skills through mentoring, represented by the function $\mu(\cdot)$, which depends on the proportion of managers which match an employee's type,[11] though not the ability of these managers.[12] This mentoring function is assumed to be increasing and continuous. The overall (lifetime) contribution to the firm's profit from promoting an applicant, which we will refer to as the applicant's "surplus," is the following function of her type, her index and the composition of the firm when she is promoted:[13]

$$\begin{aligned} s_A(\theta, m) &= x(\theta) + \mu(m) \text{ and} \\ s_B(\theta, m) &= x(\theta) + \mu(1-m). \end{aligned}$$

We denote by $z^t \in [0, r]$ the measure of the new managers in period t who are of type A; $r - z^t$ are then of type B.

The firm's per-period profit function is the total surplus generated by its new managers:

$$\pi(m, z) = \int_0^z s_A(\theta, m) d\theta + \int_0^{r-z} s_B(\theta, m) d\theta.$$

The firm seeks to maximize the discounted sum of its per-period profits by choosing a sequence of promotion policies $(z^1, z^2, ...)$ to solve the following maximization problem:

$$\begin{aligned} V(m) = \max_{(z^1, z^2, ...)} & \sum_{t=0}^{\infty} \delta^t \pi(m^t, z^t) \\ & \text{s.t. } m^{t+1} = (1-r)m^t + z^t, \\ & z^t \in [0, r] \text{ and } m^0 = m \end{aligned} \tag{1}$$

Given a value function, we can define the optimal policy correspondence by

$$z^*(m) = \{z \in [0, r] | V(m) = \pi(m, z) + \delta V(z + (1-r)m))\}$$

We pause to establish that the above optimization problem is well behaved.

[11] For our theory, it does not matter whether mentoring occurs before or after promotion.

[12] Allowing the mentoring bonus to depend on the quality of the current upper-level workers would diminish the firm's willingness to sacrifice ability of promoted workers in the current period in order to move towards more profitable levels of diversity, since less qualified workers promoted today will be less effective mentors tomorrow.

[13] These contribution functions do not include the indirect contribution from future mentoring of other workers.

Lemma 1 *The value function defined by (1) is unique and continuous. The policy correspondence z^* is compact valued and upper hemi-continuous. V and $|z^*(m) - r/2|$ are both symmetric about $m = 1/2$.*

Our model reduces to a dynamic programing problem where the state variable is the diversity of the firm's management in any period, m^t, and the optimal promotion policy depends only on the state variable. Thus, an optimal sequence of policies is given by $\{z^t\}$, where $z^t \in z^*(m^{t-1})$. We are interested in how m^t evolves over time and in how the firm biases its promotions. We define the *unbiased* (or myopic) promotion policy $z^{UB}(m)$ implicitly from the equation

$$s_A(z^{UB}, m) = s_B(r - z^{UB}, m),$$

which equalizes the contribution of the marginal new manager of each type. If the worst type A is better than the best type B ($s_A(r, m) > s_B(0, m)$) so that the marginal contributions cannot be equalized, then the firm only promotes type A's and $z^{UB}(m) = r$. Similarly, if $s_B(r, m) > s_A(0, m)$ then $z^{UB}(m) = 0$.

This myopic promotion policy ignores the dynamic consequences of promoting a given type and simply maximizes the current period payoff. We can now define the bias in the firm's optimal promotion policy as

$$b(m) = z^{UB}(m) - z^*(m).$$

A positive bias $b(m) > 0$ is then a bias in favor of type B, so that the contribution of the least qualified type A that is promoted is greater than the contribution of the least qualified type B who is promoted.[14]

Given the promotion policies z^{UB} and z^*, we can define transition functions $M^{UB}(m)$ and $M^*(m)$ which give the state achieved in period $t+1$ as a function of the period t state, m:

$$M^{UB}(m) = m(1 - r) + z^{UB}(m).$$

$$M^*(m) = \{m' | m' = (1 - r)m + z, \text{ where } z \in z^*(m)\}.$$

Our model is summarized in Figure 1. The Figure plots the supply of "final" ability, $s_A(\theta, m)$ and $s_B(\theta, m)$, when $m = .9$, for some specific functional forms. Since type A is the majority, the surplus function is higher by the amount of the mentoring differential, in this case $\mu(.9) - \mu(.1)$. When unbiased promotion policies are used, by definition, the surplus of the marginal worker is the same across the worker types, so that the firm promotes mostly type A workers. In this example, the optimal bias is positive, so that when the optimal promotion policy is used, fewer type A workers are hired, and the marginal type A worker is better than the marginal type B worker.

[14]Note that there is a one-to-one mapping from a sequence of promotion policies to both a sequence of diversity levels and to a sequence of biases. Thus while we have set up the firm's problem in terms of promotion policies, one can also think of the firm as choosing an optimal sequence of either diversity levels or biases.

2.1 Interpretations

2.1.1 Internal Promotions Versus External Hiring

We have presented our model in terms of an internal labor market (ILM) where the firm is managing the diversity of the upper of two levels in its organization. An alternative interpretation of our model is that it represents a firm where only the "upper-level" workers are employed within the firm, while the "lower-level" workers are hired from an external labor market, where ability of these workers can be observed by the firm. Then, our model applies to bias in a firm's hiring policies.

We simply need to reconsider the justification for our assumptions about the surplus received by the firm for each worker. As in the ILM interpretation, surplus in the ELM model may be justified by labor market frictions or match-specific abilities. We can assume that the firm anticipates the mentoring that will be recieved by new hires. We might also consider the case where workers of the same type are identical in terms of the value they provide to the firm, but the wage that the firm must pay to workers of a given type is increasing in the number of workers hired of each type. That is, the supply curve for each type of worker is upward sloping, so that hiring only one type of worker will increase the firm's average labor costs. This interpretation is more appropriate for labor markets in which individuals are easily substitutable in the firm's production function.

2.1.2 The Definition of Bias

Our model points out some of the subtleties in defining a promotion "bias". We think that there are at least three reasonable definitions of bias even in our simple model. Given the identical initial ability of each type, one could define bias as any deviation from equal promotion rates (i.e. as $|z(m) - r/2|$). Such a definition produces maximal "discrimination" against the minority, since majority applicants have a mentoring advantage. Alternatively, one could define the bias as any deviation from promoting the applicant who contributes the most to long-run firm profits, where the contribution is broadly defined to include not only the initial ability and human capital acquired from mentoring, but also the worker's impact on the human capital of future hires as an upper-level worker. Then a profit-maximizing firm would by definition have no bias in promotions.

We take a middle road and define bias as deviations from promoting the workers who contribute most to the firm "directly" through their own human capital, net of the worker's expected future impact on new hires through mentoring. One attractive feature of this definition is that the "direct" effect of the worker is realized no matter what promotion policy the firm uses in the future, while in contrast the future effect of the worker through mentoring itself depends on the firm's (endogenously determined) promotion policies.

Further, using our definition, bias summarizes a firm's underlying preferences towards diversity; a firm with a positive bias also sees long-run profits increasing in diversity, while a firm with a negative bias sees profits decreasing with diversity. Moreover, a firm whose profits are increasing with diversity will not only bias promotion decisions, but it will also consider the impact of a broad range of actions which might advance or hinder the careers of minority managers. For example, a firm which under our definition biases in favor of the majority may also allow a "locker room" culture that discourages minority types, while a firm that biases in favor of the minority may also hire a diversity manager.

In addition, our definition of bias is attractive in that it focuses on deviations from "type-blind" promotion policies, where the employee with greatest ability for the job is promoted ignoring issues of type. While such a distinction may not be of obvious economic importance, the question of whether firms are promoting "the most qualified person for the job" independent of skin color, gender or cultural background are at the heart of popular debate.[15]

By defining bias as the deviation from the myopic promotion policy, we emphasize the conflict between immediate profits and the long-term interests of the firm. In any case, our choice of a definition of bias is merely a matter of interpretation which does not affect the analytical results.

3 The Optimal Promotion Policy

We begin our analysis of the model by examining whether the firm's optimal promotion policy is biased in favor of minority-type workers or majority-type workers. As a building block, we first characterize the effect of diversity on the profit of a firm which lives for only one period. The one-period model highlights the trade-off between maximizing the mentoring gain for one of the applicant pools, and exploiting the scarce initial talent in both pools.

3.1 The One Period Problem

Consider a firm which lives for only one period (i.e. $\delta = 0$). The optimal promotion policy for such a firm is unbiased ($z^{UB}(m)$) since the firm does not care about the future mentoring benefits derived from promoted workers. The profit of the firm at its optimal policy is then

$$\pi^{UB}(m) = \pi(m, z^{UB}(m)).$$

[15]This argument presumes that most of the mentoring occurs before promotion. Thus, this argument does not hold for the labor market interpretation of the model developed in Section 2.1.1.

Suppressing the dependence of z^{UB} on m, we can rewrite this as follows:

$$\pi^{UB}(m) = z^{UB}\mu(m) + (r - z^{UB})\mu(1-m) + \int_0^{z^{UB}} x(\theta)d\theta + \int_0^{r-z^{UB}} x(\theta)d\theta.$$

By the envelope theorem, the effect of a change in the initial level of diversity is given by:

$$\frac{d\pi^{UB}(m)}{dm} = z^{UB}\mu'(m) - (r - z^{UB})\mu'(1-m). \tag{2}$$

Increasing the initial proportion of type A workers leads to an increase of $\mu'(m)$ in the mentoring received by the z^{UB} type A workers who will be promoted, as well as a corresponding reduction of $\mu'(1-m)$ experienced by the $r - z^{UB}$ workers who will be promoted. Thus, the sign of $d\pi^{UB}(m)/dm$ depends on the curvature of the μ function as well as the relative proportions of type A and B workers promoted. If the mentoring function is concave, then $\mu'(1-m) > \mu'(m)$, and increases in diversity (i.e. a lower m) increase the mentoring of minority employees more than they decrease the mentoring of majority hires. However, since the mentoring function is nondecreasing, majority types gain more from mentoring than minority types, and the unbiased firm hires more of the (majority) A types than the (minority) B types ($z^{UB} > r/2 > r - z^{UB}$); thus, profits are more sensitive to the mentoring of A types than B types.

As a result, the firm's profit increases with the diversity of its management only if the mentoring function is concave *and* the degree of concavity is sufficient to overcome the larger weight placed on mentoring outcomes for majority workers. We shall say that *sufficient concavity* (SCV) holds at $m > 1/2$ if

$$\mu'(m) < \frac{r - z^{UB}(m)}{z^{UB}(m)}\mu'(1-m).$$

Further, we say that *sufficient concavity holds everywhere (SCV everywhere)* holds if SCV holds for all $m > 1/2$. Similarly, we shall say that *sufficient concavity holds nowhere (SCV nowhere)* if μ fails to be SCV for each $m \geq 1/2$. Note that convexity of μ implies nowhere SCV.

By definition, if SCV holds everywhere then the profits from unbiased promotions increase with diversity. That is, $\frac{d\pi^{UB}(m)}{dm} < 0$, for $m > \frac{1}{2}$. Similarly, if SCV holds nowhere, then profits from unbiased promotions are decreasing with diversity. That is, $\frac{d\pi^{UB}(m)}{dm} > 0$, for $m > \frac{1}{2}$.

Now consider the optimal promotion policy of a firm that lives for two periods. In the second (and final) period, the firm does not care about future mentoring and unbiased promotions are optimal; thus, equation (2) characterizes the firm's preferences over the level of last-period diversity. Consequently, if SCV holds everywhere, the most profitable level of diversity at the beginning of the final period is full diversity

(i.e. $m = 1/2$). This leads the firm to bias its first period promotions in favor of the minority. Conversely, SCV holds nowhere, the best state at the beginning of the final period is homogeneity (i.e. $m \in \{0, 1\}$). This leads the firm to bias its first period promotion in favor of the majority type.

3.2 Characterizing Sufficient Concavity

We have shown that condition SCV determines the first-period promotion bias for a firm which lives two periods. Throughout the rest of the paper, we will show that condition SCV is useful in characterizing promotion biases and long-run dynamics for infinitely-lived firms. Thus, we pause to further characterize and explore this condition.

3.2.1 The Mentoring Function

The condition SCV is a restriction on the shape of the mentoring function relative to other parameters of the model. In many settings, we expect that the mentoring function should exhibit diminishing returns. In its most literal sense, mentoring refers to a voluntary relationship between an experienced employee and a new hire, generally of the same type. An increase in the proportion of managers of a given type may allow lower level employees who would otherwise be unmatched to find a mentor; it also may allow them to find a more appropriate mentor than they otherwise would have. While the first minority-type upper-level workers yield high returns by providing mentoring to those who would not otherwise have received it, as more minorities are promoted, additional upper-level minorities simply provide somewhat better mentoring possibilities to those who would have been mentored anyway.

We also allow for the mentoring function to exhibit constant or increasing returns. The last few steps toward homogeneity may be the most valuable in terms of mentoring if homogeneity in language and culture greatly increases the efficiency of communication and information sharing. For example, the advantages of a highly cooperative management culture might be significantly compromised by the presence of even a few managers who are not cooperative types. In such instances, the SCV will hold nowhere.

A third possibility for the shape of the mentoring function is what we call a *critical mass* mentoring function, which is first convex and then concave. This mentoring function incorporates the idea that $\mu(m)$ increases slowly at first due to the relative ineffectiveness of mentors who are themselves in the minority. However, the returns to additional mentors increase rapidly in the neighborhood of a "critical mass" of mentors. Once the critical mass is achieved, diminishing returns set in. For such mentoring functions, we expect that SCV will hold in some regions but not others.

3.2.2 The Exogenous Parameters

In addition to the shape of the mentoring function, several other exogenous parameters affect whether condition SCV holds. First, we consider the retirement rate, r. This represents the rate of turnover relative to the size of the firm. Industries with low turnover at the senior level relative to the size of the firm include law, academics, businesses organized by partnership, and high levels of large bureacracies; industries with high turnover include many high-technology industries, as well as intermediate levels of hierarchies.

We will further introduce a parameter, α , which describes the importance of mentoring relative to initial ability. For any mentoring function $\widehat{\mu}$, we can define $\mu(m) = \alpha\widehat{\mu}(m)$, and explore the effects of varying α. Industries where mentoring is likely to be important relative to initial ability include services, law, academics, industries where networking and information sharing play an important role, and industries where apprenticeships are important components of training. In contrast, mentoring will be less important in industries where jobs are well-defined, require few specialized skills, individual production, and involve little information sharing.

Finally, we will introduce a parameter γ, which describes the scarcity of initial ability. An increase in γ makes x steeper, so that for $z > \frac{r}{2}$, $x(z;\gamma) - x(r-z;\gamma)$ is nondecreasing in γ. Thus, for any given promotion rule which deviates from equal proportions in promotion, then the higher is γ, the larger is the cost in terms of initial ability of the unbalanced promotion decision. Observe that our assumptions are satisfied for a linear initial ability function $x(\theta;\gamma) = \kappa - \gamma\theta$. Industries where scarcity is important include industries which require specialized skills and experience, such as high-level management, or industries where "stars" are important, such as academics.

The following proposition summarizes the effects of these parameters on the condition SCV, which holds if $\mu'(m) - \frac{r-z^{UB}(m)}{z^{UB}(m)}\mu'(1-m) \leq 0$.

Proposition 2 *For $m > 1/2$, the expression $\mu'(m) - \frac{r-z^{UB}(m)}{z^{UB}(m)}\mu'(1-m)$ is:*
(i) nonincreasing in γ;
(ii) nondecreasing in α;
(iii) nonincreasing (nondecreasing) in r whenever $\frac{d}{dr}z^{UB}(m;r,\alpha,\gamma)\cdot\frac{r}{z} \leq (\geqslant) 1$.

Proposition 2 demonstrates several natural properties of the firm's optimal policy. First, an increase in scarcity (as parametrized by γ) increases the cost of an asymmetric promotion policy, thereby increasing the benefits of diversity so that SCV is more likely to hold. Second, an increase in the importance of mentoring relative to initial ability (as parametrized by α) increases z^{UB}, since the majority type gains a larger advantage, and thus makes SCV less likely to hold. When α becomes sufficiently large, the firm's optimal policy will only select majority types (SCV will hold nowhere); for α sufficiently close to zero, unbiased promotion selects almost the same number of majority and minority types, and then concavity and SCV are (approximately) equivalent.

Finally, an increase in r requires the firm to promote more workers, which changes the proportion of type A workers hired. If the elasticity of z^{UB} with respect to r is less than 1, the *proportion* of type A workers falls when r increases, and so condition SCV is more likely to hold. When the initial ability function takes the specific form $x(\theta) = \kappa - \gamma\theta$, the elasticity restriction holds, and an increase in r decreases (SCV).

3.3 The Infinite Horizon Problem

In this section, we show that our characterization of firm's bias when it lives only two periods extends to the case where the firm faces an infinite horizon, so long as the firm's one-period value function, $\pi^{UB}(m)$, is monotonic throughout the region $m \geqslant \frac{1}{2}$.

Proposition 3 *Consider $m > 1/2$. (i) Suppose that SCV holds everywhere. Then the value of the firm is increasing with the level of diversity ($dV/dm < 0$) and the optimal promotion policy is biased in favor of the minority ($b(m) > 0$). (ii) Now suppose that SCV holds nowhere. Then the value of the firm is increasing as the upper level becomes more homogeneous ($dV/dm > 0$) and the optimal promotion policy is biased in favor of the majority ($b(m) < 0$).*

The proof of this Proposition uses induction on the number of periods to show that if the one-period value function is nonincreasing, then the infinite horizon value function must be as well. The key feature of our model for this result is the fact that $\frac{\partial^2}{\partial m \partial z}\pi(m,z) \geqslant 0$: the higher the initial proportion of type A workers, the higher are the returns to increasing the number of type A workers promoted. Consider the case where SCV holds everywhere, and suppose that the future value of the firm starting in period 2 is nonincreasing in m. This will lead a forward-looking firm in period 1 to choose more diversity than it would using an unbiased promotion policy. Since SCV everywhere implies that an increase in diversity improves per-period profits when *unbiased* promotion policies are in place, increasing the initial level of diversity will be yet more beneficial when the *optimal* promotion policies are used, since $\frac{\partial^2}{\partial m \partial z}\pi(m,z) \geqslant 0$.

Proposition 3 shows that, just as in the single-period problem, the firm may prefer either homogeneity or diversity even if attention is limited to concave mentoring functions. If SCV holds everywhere, then the greater impact of additional minority managers on future mentoring always offsets the greater importance of effective mentoring of the majority pool. Then the firm benefits from increases in diversity (i.e. m closer to $1/2$), which translates into a promotion bias in favor of the minority. Conversely, if SCV holds nowhere, then the greater importance of effective majority mentoring always dominates, the firm prefers a more homogeneous management (i.e. m closer to 0 or to 1) and the promotion bias is in favor of the majority.

There are, of course, parameter values for which SCV holds in some regions but not others. For these mentoring functions it is possible that the firm biases in favor

of the majority for some values of m and in favor of the minority for others. Consider the following example: $\mu(m) = m^{.11}$, $x(\theta) = 1 - \theta$, and $r = .3$. Figure 2 plots the one-period value function, $\pi^{UB}(m)$. Note that profits increase with diversity on $[\frac{1}{2}, .927]$ but not on $[.927, 1]$. Figure 3 illustrates the unbiased promotion rate (z^{UB}) for type A applicants as well as the optimal promotion rates for $\delta = .3$ and $\delta = .95$. The 45^o line is the retirement rate, since we assume a uniform retirement rate in each period. The bias in favor of type B workers, $b(m; \delta)$, is computed as the difference between the unbiased promotion rate ($\delta = 0$) for type A 's and the biased promotion rates.

Figure 3 shows that for $\delta = .95$, the firm always biases in favor of the minority; despite the fact that the one-period value function is increasing near 1, the firm finds it worthwhile to bias promotions towards full diversity, which maximizes the one-period value function. In contrast, if $\delta = .3$, the firm's optimal policy is to bias in favor of the majority in the region $[.93, .96]$, and in favor of the minority elsewhere. Hence, we show that even within a single firm, attitudes towards diversity will change with the initial conditions. We will further discuss the dynamics of this example in the next section.

4 The Dynamics of Diversity

We now turn to a characterization of the evolution of diversity at our firm. Our model can be used to analyze the dynamic path followed by firms from any initial condition to a steady state, and we illustrate these dynamics with numerical examples. However, most of our emphasis is on steady state levels of diversity. We provide conditions under which diversity and homogeneity are steady states, and further we study conditions under which there are multiple steady states. We then conduct comparative statics analysis.

4.1 Convergence to a Steady State

A steady state level of diversity is defined as a fixed point in the optimal transition correspondence $m_s \in M^*(m_s)$, so that it is optimal for a firm with diversity m_s to maintain that level. In terms of the optimal promotion policy, a steady state satisfies $rm_s \in z^*(m_s)$. That is, at a steady state, the measure of retiring A managers exactly equals the measure of type A's that are promoted. Similarly, a steady state with unbiased promotions satisfies $m_s = M^{UB}(m_s)$ and $rm_s = z^{UB}(m_s)$. Figure 3 graphs the promotion policy against the (uniform) retirement function; each crossing of the two functions represents a steady state. In Figure 3, the set of steady states when $\delta = .3$ is given by $\{0, .067, .5, .936, 1\}$.

The following proposition proves that the level of diversity always converges to a steady state.

Lemma 4 *(i) $M^{UB}(\cdot)$ is a single-valued, nondecreasing continuous function.*
(ii) $M^{*}(\cdot)$ is a single-valued, nondecreasing and continuous function except for possible upward jumps, where the correspondence will contain both the high and low values.
(iii) For $m > \frac{1}{2}$, $M^{*}(m) > \frac{1}{2}$ and $M^{UB}(m) > \frac{1}{2}$.
(iv) With either biased or unbiased promotions, the diversity of the firm converges to a steady state in $[\frac{1}{2}, 1]$ for $m^0 > \frac{1}{2}$.

Lemma 4 establishes that, under either unbiased or optimal promotion policies, the more type A's in management today, the more there will be tomorrow. This result (similar to Theorem 3) follows from the fact that $\frac{\partial^2}{\partial m^t \partial m^{t+1}} \pi(m^t, m^{t+1} - rm^t) \geq 0$: in terms of today's profits, the marginal return to increasing the future proportion of type A workers is nondecreasing in the proportion of type A workers in the firm today.

Further, Lemma 4 shows that in our model, the majority type is fixed. That is, if $m^0 > \frac{1}{2}$, then $m^t > \frac{1}{2}$ for all t. A transition function that is nondecreasing and continuous (but for upward jumps) is sufficient to establish convergence to a steady state.

Further, with type based mentoring and $m^t > \frac{1}{2}$, the immediate returns from moving to any $m^{t+1} > \frac{1}{2}$ are greater than the returns from moving to $1 - m^{1+1}$. This is sufficient to establish the result since by the symmetry of the value function m^{t+1} and $1 - m^{t+1}$ have the same continuation value (i.e. $V(m^{t+1}) = V(1 - m^{t+1})$).

This is the first of several propositions that explores the extent to which history matters in our model. Because we have a nondecreasing transition function, the majority type is never pushed into the minority. We now consider whether the initial conditions dominate in the long run (as opposed to the alternative, where the firm moves towards full diversity ($m = \frac{1}{2}$)).

4.2 The Set of Stable Steady States

We distinguish between two types of steady states. Consider Figure 3 again, and focus on the case where $\delta = .3$. There is an important difference between the steady state at .936, and the ones at 1/2 and 1. Diversity only converges to .936 if $m^0 = .936$. If the firm starts at $m^0 = .936 + \varepsilon$, where ε is small, then diversity converges to $m = 1$ for $\varepsilon > 0$ and to $m = 1/2$ for $\varepsilon < 0$. We refer to points such as .936 in Figure 3 as *critical mass points,* since the long-run outcome of the firm is determined by whether the initial conditions achieve a level of diversity greater than the critical mass. We shall refer to a steady state which is always reached from a neighborhood surrounding it as a *stable steady state.* Formally:

Definition 1 *A level of diversity m is* **stable** *(***a critical mass point***) for some transition function M if there exists a $v > 0$ such that, for all $v > \varepsilon > 0$ such that $M(m + \varepsilon)$ and $M(m - \varepsilon)$ are single-valued, $m + \varepsilon \gtrless (<) M(m + \varepsilon)$ and $m - \varepsilon \leq (>) M(m - \varepsilon)$.*

Let S^* be the set of stable steady states under the optimal promotion policy and let S^{UB} be the set of stable steady states with myopic promotions. We will refer the following notion of symmetry when analyzing stable steady states: S is symmetric if $m \in S \Rightarrow 1 - m \in S$. Since we have assumed that the two types are symmetric, our firm's policies and steady states will be symmetric as well. We use this fact to simplify our exposition.

We start by characterizing S^{UB}. This is of interest because a variety of factors (such as legal restrictions or agency problems) may lead some organizations to pursue myopic promotion, and further because it helps us to isolate the effect of the optimal bias on the set of steady states.

4.2.1 Characterizing the Stable Steady States with Unbiased Promotions

The following Proposition characterizes stable steady states when the firm uses an unbiased, or myopic, promotion policy:

Proposition 5 *(i)* $\frac{1}{2} \in S^{UB}$ *iff* $\mu'(\frac{1}{2}) < -rx'(\frac{r}{2})$.
(ii) For any finite, symmetric set of points $X \subset [0,1]$, *there exists a nonincreasing initial ability function* $x(\cdot)$ *and a nondecreasing mentoring function* μ *such that* $S^{UB} = X$.

Full diversity ($m = \frac{1}{2}$) is stable when the effect of increasing m on the surplus differential between workers that arises due to type based mentoring ($\mu'(\frac{1}{2})$), is small relative to the effect on initial ability ($rx'(\frac{r}{2})$).[16]

Part (ii) derives conditions for multiple stable steady states, showing that S^{UB} can take (almost) any form long as it is symmetric. This extreme result uses very particular functional forms. For many functional forms we expect that there will be no more than three steady states. Consider the following example:

Example 1 *Suppose the initial ability function is linear,* $x(\theta) = \kappa - \gamma\theta$. *If* $\mu'''(m) < 0$ $\forall m$, *then* $S^{UB} = \{m, 1-m\}$ *for some* $m \geq \frac{1}{2}$. *If* $\mu'''(m) > 0$ $\forall m$, *then* $S^{UB} \subset \{0, \frac{1}{2}, 1\}$.

When the initial ability function is linear and $\mu'''(m) > 0$, the transition function M^{UB} is convex; thus, firms who simply hire the most "able" workers will eventually attain either full diversity or full homogeneity, but no intermediate outcome is stable. In contrast, if $\mu'''(m) < 0$, the transition function M^{UB} is concave, and an intermediate outcome is possible. Full diversity might never be attained; but if it is attained, it is the unique long-run stable steady state. Figure 3 graphs promotion and retirement functions for a linear initial ability function and a mentoring function in the class $\mu(m) = \alpha m^{\rho}$; note that for this functional form, μ is nondecreasing and concave iff $\mu''' > 0$. Figure 3 shows that $S^{UB} = \{0, \frac{1}{2}, 1\}$.

[16]Here, as elsewhere, the scarcity of initial ability depends on the product of the retirement rate and the derivative of the initial ability function x . The retirement rate enters in because it determins the rate at which the firms dips into its applicant pools.

4.2.2 Stable Steady States with Optimal Promotion Policies

We now explore the effect of the optimal bias on the evolution of diversity. We can write

$$M^*(m) = M^{UB}(m) + b(m),$$

illustrating that the effect of the optimal bias is to shift the unbiased transition function. To further analyze this effect, restrict attention to the interval $[\frac{1}{2}, 1]$. If the bias has an unambiguous sign, then the transition function shifts up when the bias is positive and down when it is negative. Such monotonic shifts in the transition function produce monotonic shifts in the set of stable steady states (Milgrom and Roberts, 1994). Thus, when SCV holds everywhere or SCV holds nowhere, the set of stable steady states using optimal promotions is comparable to the set of stable steady states attained by a myopic firm.

Definition 2 *Consider two sets of stable steady states, X_1 and X_2. We say that X_1 is* **more diverse** *than X_2 if*

$$\min\{x|x \in X_2, x \geq \frac{1}{2}\} \geq \min\{x|x \in X_1, x \geq \frac{1}{2}\}$$
$$\max\{x|x \in X_2, x \geq \frac{1}{2}\} \geq \max\{x|x \in X_1, x \geq \frac{1}{2}\}$$

Building on Proposition 5, which characterizes the set of stable steady states under unbiased promotion policies S^{UB}, the next proposition characterizes S^* relative to S^{UB}.

Proposition 6 *(i) If SCV holds everywhere, then S^* is more diverse than S^{UB}. If SCV holds nowhere, then S^{UB} is more diverse than S^*.*
(ii) If SCV holds nowhere, then $\frac{1}{2}$ is never a stable steady state.

Part (i) shows that implementing an optimal bias which is in favor of the minority (as when SCV holds everywhere) increases long-run diversity, while a bias in favor of the majority (as when SCV holds nowhere) decreases long-run diversity. As a result, the effect of public policies which seek to force firms to hire "the best person for the job" can have ambiguous effects on diversity. The result has implications for the status of full diversity and homogeneity. In particular $\frac{1}{2} \in S^*$ whenever $\frac{1}{2} \in S^{UB}$; $S^* = \{\frac{1}{2}\}$ whenever $S^{UB} = \{\frac{1}{2}\}$; and $S^* = \{0, 1\}$ whenever $S^{UB} = \{0, 1\}$. Part (ii) establishes that full diversity can never be a stable steady state when SCV holds nowhere.

When full diversity is the unique outcome, then history does not matter in the long run. If there is a unique stable steady state level of diversity other than $\frac{1}{2}$ (i.e. $S^* = \{m, 1-m\}$), then the only effect of initial conditions is to determine which type is the majority. In such a case, our model produces a phenomenon which resembles the "glass ceiling" described in the popular press. In particular, a firm which starts

with homogeneous management will diversify initially. However, at some point the progress of minorities into management stalls. There is an "invisible barrier" (type-based mentoring) which keeps minority representation in management from mirroring the diversity of the lower level of the organization.

Figure 3 gives an example where there are multiple steady states when $\delta = .3$. Further, it can be shown that even if SCV holds everywhere (which is not the case in Figure 3), there may exist multiple stable steady states $m_s^H > m_s^L \geq 1/2$.[17] In such a case, initial conditions may not just determine which type is the majority (as when $S^* = \{m, 1-m\}$), they may also determine the level of long-run diversity. As in Figure 3, there can be a critical mass at some point $\bar{m}$. Then the firm gets "stuck" and does not reach the stable steady state with highest diversity if $m^0 > \bar{m}$. One implication of these dynamics is that short-lived pressure on the firm to diversify (e.g. from government regulation or the popular press) can have long-lived effects on the level of diversity. Suppose the firm starts at $m^0 = m_s^H$ but at some point is subjected to pressure to diversify further. If the pressure shifts the level of diversity above the critical level in (m_s^L, m_s^H), then the gains of the minorities will be self-reinforcing. The long-run level of diversity will be m_s^L even if the pressure is removed.

When SCV holds everywhere and there are multiple steady states, Proposition 3 ranks the steady states in terms of profits: the less diverse steady state m_s^H represents an "inertial trap" with lower profits. Thus, we see that firms employing the same technologies and facing identical worker pools can have different levels of profitability depending on historical conditions which affect their initial levels of diversity. While the less profitable firm could imitate the organization of the other firm, the existence of multiple steady states implies that the cost of the transition is more than the less diverse firm is willing to bear.

When SCV holds for some levels of diversity but not others, the bias can have an ambiguous effect on diversity. Return to Figures 2 and 3, where $S^{UB} = \{0, \frac{1}{2}, 1\}$, but per-period profits are maximized at $m = \frac{1}{2}$. For relatively short-sighted firms ($\delta = .3$), moving to full diversity from a homogeneous inital condition incurs prohibitive transition costs, and thus a firm which starts at $m^0 \in [.936, .96]$ will reach full homogeneity faster than with unbiased promotions. On the other hand, a far-sighted firm ($\delta = .95$) is willing to incur the transition costs to get to the best per-period level of diversity: the far-sighted firm always biases in favor of the minority, and the firm always converges to full diversity, $S^* = \{\frac{1}{2}\}$.

A second example is illustrated in Figures 4 and 5. Figure 4 illustrates a "critical mass" mentoring function, where mentoring is relatively ineffective until the minorities reach 20% of the firm. Unbiased promotions (or optimal promotions by a short-sighted firm) lead to stable steady states at $S^{UB} = \{.1, .5, .9\}$. However, a far-sighted firm ($\delta = .95$) recognizes that it receives maximal per-period profits at full diversity, and

[17] For example, if $\mu(m) = .1m^{.1}$, $x(\theta) = 1 - \theta$, $r = .05$, an example can be constructed with the highest stable state being $m = 1$.

thus it biases promotions to achieve $S^* = \{\frac{1}{2}\}$. Figure 5 graphs the proportion of type A managers in each period for a short-sighted firm, given a variety of different initial conditions. Again, short-sighted firms appear to have a "glass ceiling," because starting from homogeneity, minorities make some initial progress into the firm; however, they never reach a level of representation above 10%.

4.3 Comparative Statics

Next, we turn to the comparative statics on the set of steady states.[18] We introduced the parameters of interest in Section 3.2.2; our analysis in this section derives predictions about how the level of diversity varies across different economic environments.

The effect of any parameter on the long-run steady states of profit-maximizing firms can be decomposed into two parts, the effect on the unbiased dynamics and the effect on the optimal bias. We start by characterizing the first effect.

Proposition 7 *The diversity of S^{UB} is decreasing in α, increasing in γ , and is independent of δ. If $\partial z^{UB}/\partial r < 1/2$, then diversity is increasing in r.*

Proof: Recall that $m \in S^{UB}$ satisfies $z^{UB}(m) - rm = 0$. Consider $m \geq \frac{1}{2}$. Then the diversity of S^{UB} is increasing (decreasing) in a parameter if $z^{UB}(m) - rm$ is decreasing (increasing) in that parameter. It is straightforward to show that $\partial z^{UB}/\partial \alpha > 0$, $\partial z^{UB}/\partial \gamma < 0$ and $\partial z^{UB}/\partial \delta = 0$. Finally, if $\partial z^{UB}/\partial r < 1/2$, then $\partial z^{UB}/\partial r - m < 0$ for all $m \geqslant \frac{1}{2}$. □

To interpret Proposition 7, observe that as mentoring becomes more important, majority candidates become more attractive and their promotion rates increase. As a result, diversity falls with α. Conversely, as initial ability becomes scarcer (high γ), mentoring has less impact on promotion rates and diversity increases. Finally, increases in the retirement rate has two effects. First, because there are more majority type managers, they have more to lose from rapid retirement, and hence increases in r can favor diversity. Second (and similar to our analysis of sufficient concavity), as r increases the firm dips deeper into its applicant pool, and the curvature of the initial ability function, $x(\theta)$, determines whether promoting more workers increases or decreases the cost of asymmetric promotion decisions.[19] If r increases minority promotions faster than majority promotions using the unbiased policy (as is true if $x(\theta) = \kappa - \gamma\theta$), then increases in r increase diversity.

[18] As will be discussed further below, our comparative statics analysis is restricted to a local approach, where we consider the effect of small changes in parameters on any given stable steady state.

[19] Specifically, with a convex initial ability function, the ability of the minorities falls off faster than that of the majority as more of each are promoted. Hence, an increase in r can lead to a decrease in diversity. Conversely, if x is concave, then the ability of the majority types falls off faster, which favors diversity.

Now we turn to analyze how the set of stable steady states changes with exogenous parameters when optimal promotion policies are used. A natural first approach to this problem would be to derive comparative statics results on the optimal bias, and use these results to analyze changes in steady states. Since the bias increases exactly when the value function gets steeper, this approach is equivalent to deriving comparative statics on the slope of the value function. In Proposition 3, we showed that when the one-period value function is nonincreasing (nondecreasing) in m, so is the infinite horizon value function. Unfortunately, similar connections do not in general hold for other properties of the value function: making the one-period value function steeper everywhere does not necessarily make the infinite-horizon value function steeper everywhere, without additional restrictions on the curvature of the one-period payoff function.

Instead, we take a more "local" approach. Rather than analyze changes in the bias at arbitrary initial conditions, we focus only on comparative statics for states m_s which are steady states. Our analysis makes use of the fact that $z^*(m_s) = rm_s$ to compare the magnitudes of the potentially competing effects of exogenous paramters on present versus future profits. Despite the fact that our approach is somewhat more restrictive, we still find that several of our results from Section 3.2.2 must be qualified in the fully dynamic problem.

We begin by examining the First Order Conditions for the unique optimal promotion rule which must be satisfied in a stable steady state:

Lemma 8 *For all $m_s \neq \frac{1}{2}$ which are stable steady states, there is a unique optimal promotion policy, $z^*(m_s)$. An interior steady state $m_s \notin \{0, \frac{1}{2}, 1\}$ satisfies the following first order condition*

$$FOC_s(m_s) = \pi_2(m_s, rm_s) + \frac{\delta}{1-(1-r)\delta}\pi_1(m_s, rm_s) = 0.$$

The First Order Condition illustrates that the return to promoting more majority type workers today is balanced against to the (weighted) return of beginning a new period with a higher level of diversity. Given uniqueness of the optimal promotion policy, the implicit function theorem together with differentiability of the objective function implies that a small increase in $FOC_s(m_s)$ must lead to an increase in the stable steady state. We exploit this fact in our comparative statics analysis.

Proposition 9 *Each stable steady state $m_s \gtrless \frac{1}{2}$ is:*
(i) Nonincreasing in response to small increase in γ.
(ii) When SCV holds everywhere (nowhere), nonincreasing (nondecreasing) in response to a small increase in δ.
(iii) When SCV holds nowhere, nondecreasing in response to a small increase in α.
(iv) When SCV holds everywhere and $x(\theta) = \kappa - \gamma\theta$, nonincreasing in response to a small increase in r.

Proof: (i) Applying the implicit function theorem, since by assumption m_s is optimal, it suffices to check that

$$\frac{\partial FOC_s(m;\gamma)}{\partial \gamma} = \frac{\partial [x(rm;\gamma) - x((1-r)m;\gamma)]}{\partial \gamma} \geq 0$$

for $m > \frac{1}{2}$, which holds by definition. (ii) See appendix. (iii) and (iv): See text below. □

Proposition 9 shows that the faster initial ability declines with the proportion of a given type promoted, the more diverse the upper level will be in steady state. Note that the direct effect of a shift in γ—a greater willingness to promote minorities—is reinforced by the effect of γ on the optimal bias. As the firm promotes more minorities, it cares more about their mentoring and hence it is more interested increased diversity.

The parameter δ has a simple effect in our model. As the future looms larger, the optimal bias becomes larger, since the firm cares more about the effectiveness of future mentoring. Hence, the effect on the dynamics depends on the direction of the optimal bias. If, for example, the bias is for the minority, then diversity increases. Further, as usual, if δ is close enough to 1, there is a unique long-run level of diversity.

This result identifies the nature of the inefficiency which arises if the agents responsible for promotions were to discount the future more than the firm's owners. In practice, long-term compensation schemes are rarely used for agents in the middle levels of a firm's hierarchy, and workers take into account the fact that they will leave the firm (or at least their current job) with positive probability in the future. Such a divergence of discount factors would lead to inefficiencies, according to our analysis, and firms whose mentoring functions are globally SCV might move towards diversity too slowly. In such a case, firms have an incentive to create internal rules or policies about hiring and promotion which help correct the agency problem. This is consistent with the existence of internal affirmative action policies within many large firms.

Furthermore, since such agency problems may exist even in top management, shareholders may wish to constrain the policy of top management on diversity. The ability of shareholders to dictate diversity policy has been a topic of recent debate by the Securities and Exchange Commission (SEC).[20] In 1993, the SEC reversed a long-standing policy by allowing Cracker Barrel Old Country Store to exclude a shareholder resolution which prohibited job discrimination on the basis of sexual orientation, arguing that diversity policy fell within the range of "ordinary business matters" which cannot by SEC rules be dictated by shareholder resolutions. Recently, shareholder interest groups have questioned the policy change. Our analysis suggests that shareholders may indeed have an interest in directing company policies towards diversity, even beyond the more widely cited issues of legal liability.

[20]Wallman, Steven, "Equality Is More Than 'Ordinary Business,"' *New York Times* Page 3, 12:3, March 30, 1997.

The comparative statics on S^{UB} for α and r do not extend to the case with optimal promotions. Consider first the effect of α, where $\mu(m) = \alpha\hat{\mu}(m)$. It might seem that when mentoring becomes more important, the firm would always want to increase homogeneity in order to improve the mentoring of the majority candidates. While this intuition is correct for large enough α, there is a competing effect for intermediate levels. This can be seen in reference to the first order conditions:

$$\frac{\partial}{\partial\alpha}FOC_s(m;\alpha) = \hat{\mu}(m) - \hat{\mu}(1-m) + \frac{r\delta}{1-(1-r)\delta}(m\hat{\mu}'(m) - (1-m)\hat{\mu}'(1-m)).$$

The importance of mentoring has two effects on steady states. First, it increases the mentoring advantage of majority candidates, which directly increases their promotion rates (as reflected by the term $\hat{\mu}(m) - \hat{\mu}(1-m) > 0$), a force which favors increased homogeneity. Second, as mentoring becomes more important, the firm becomes more willing to bias promotions to optimize mentoring. Since this bias may be for or against the minority, the effect of a change in the importance of mentoring is in general ambiguous for intermediate levels of α.

However, we can make additional progress in the special case where the mentoring function is nowhere SCV, as shown in Proposition 9. To see this, observe that by Proposition 3 and optimality of the stable promotion policy, SCV nowhere implies that $rm_s > z^{UB}(m_s)$. But then, the second term in $\frac{\partial}{\partial\alpha}FOC_s(m;\alpha)$ must be greater than $\frac{\partial}{\partial m}\pi^{UB}(m)$, which in turn is positive by SCV nowhere.

Now consider the rate of turnover. We obtained unambiguous comparative statics on S^{UB} for r when $x(\theta) = \kappa - \gamma\theta$. Restricting ourselves again to the linear case we find that,

$$\frac{\partial}{\partial r}FOC_s(m;r) = -(2m-1)\gamma + \frac{\delta(1-\delta)}{(1-(1-r)\delta)^2}(m\mu'(m) - (1-m)\mu'(1-m)).$$

As with α, there are two effects. There is direct effect, captured in the term $-(2m-1)\gamma$, which favors diversity. As the retirement rate goes up the firm must dig deeper into its applicant pool, and the importance of mentoring relative to initial ability falls. However, there is also an indirect effect: a firm is more willing to bias promotions as r increases since current promotions have an immediate and large effect on diversity. If the bias is opposed to the minority, the indirect effect opposes the direct effect.

However, if SCV holds everywhere, we are able to make a prediction. To see this, observe that by Proposition 3 and optimality of the stable promotion policy, SCV everywhere implies that $rm_s < z^{UB}(m_s)$. But then, the second term in $\frac{\partial}{\partial r}FOC_s(m;r)$ must be less than $\frac{1}{r}\frac{\partial}{\partial m}\pi^{UB}(m)$, which in turn is negative by SCV everywhere.

There are other potential comparative statics as well: for example, if the initial ability function for one type of worker increased relative to the other, the steady states would entail more workers of that type. If such asymmetries were important, far sighted firms might be willing to change their composition so that the better type is in the majority.

5 Market Equilibrium and Segregation

In our basic model, we have assumed that the entry level of our firm is fully integrated and that there are no labor market interactions between firms. In this section, we relax some of these assumptions (although for simplicity we consider only a static model). We are motivated by the question of whether, in a world of type-based mentoring, the existence of competitive labor markets leads to segregation, where some firms only hire A-types and some firms only hire B types. Since type A individuals will be more productive in firms with more type A workers in the upper level, type-based mentoring does indeed create a force in favor of type-based specialization by firms in the same labor market.

Of course, in practice, we do sometimes observe segregated outcomes. For example, in the U.S., several industries such as law and banking historically included some firms with a large Jewish representation. Construction companies in some U.S. cities are dominated by a particular ethnic group. Some academic disciplines have large representations from particular nationalities, and gender integration is surprisingly heterogeneous across disciplines, even within the sciences. In other examples, the software company T/Maker in Cupertino, CA was founded by a woman and the firm remained over two-thirds women after 15 years in business; as discussed in the introduction, the fashion company Esprit de Corps has approximately two-thirds women in management. However, more often firms are not strongly segregated.

How do we reconcile the lack of strongly segregated outcomes with the existence of type-based mentoring? First, there are a host of significant costs to homogeneity of a firm's lower levels. For example, it is rare that firms are truly identical from the perspective of workers. Their locations, cultures, and the skills required will differ; it may be efficient for a diverse group of workers to remain in the same firm, despite mentoring disadvantages.[21]

Instead of considering such direct costs of homogeneity, however, this section develops a static version of our basic model which incorporates two firms and a labor market for workers.[22] Not surprisingly, we find that type-based mentoring does encourage firms to concentrate their lower level hiring on one type. There is a benefit to concentrating employees of each type at the firms that have a comparative advantage in mentoring them. However, we show that such segregation need not be complete. As a firm hires more employees of the majority type, the probability that any given majority-type worker makes it to the upper level falls, and hence so does the marginal benefit of hiring majority-type workers. If differences in diversity are not too great, we show that

[21] Additional costs to homogeneity arise from government and civil litigation for biased hiring practices, social pressure on firms to mirror the composition of their local labor markets, the fact that different types may be better suited for different jobs, and worker preferences.

[22] Our model still fails to endogenize the labor market participation decisions and human capital acquisition decisions, as well as the initial conditions of existing firms, are taken to be exogenous.

there can be an interior equilibrium in which all firms hire some of both types. As with any scarce input with decreasing marginal productivity, it can be efficient to spread majority-type employees across firms.

5.1 A Two-Firm Model

Unless otherwise stated, we retain the earlier assumptions from our single-firm model. The main change is the addition of a second firm that competes for entry level employees in a competitive labor market. We simplify the analysis by limiting ourselves to a static model, where each firm makes hiring and promotion decisions only once, and by assuming that the initial ability function is linear.

Formally, there are two firms indexed by $i = 1, 2$. Each firm hires a unit measure of lower level employees at the beginning of the period. Let q_i be the proportion of lower level employees at firm i of type A and let w be the wage *differential* between A's and B's. In the labor market there is a unit measure of each type of employee, and the labor market must clear. That is, $q_1(w) + q_2(w) = 1$, where $q_i(w)$ is firm i's demand for entry level A's at a given wage differential. More generally, we could include an upward-sloping supply curve for both types of labor.

The initial ability for management in the population of both A's and B's is given by $X(\theta) = 1 - \gamma\theta$ for $\theta \in [0, 1]$. Firms and workers observe initial ability only after hiring (but before promotion) and hence each firm attracts a random selection of workers from each population. The supply of initial ability of each type of entry-level worker at a given firm then depends on how many of each type it hires. In particular, it must be that the initial ability function for type A workers at firm i is $x^A(\theta, q_i) = 1 - \gamma\theta/q_i$ while for type B workers it is $x^B(\theta, q_i) = 1 - \gamma\theta/(1 - q_i)$.

For firm i, let z_i denote the proportion of workers promoted who are of type A; let m_i denote the proportion of managers of type A; and let μ_i denote the mentoring function. The profit of firm i is the is given as follows:

$$\begin{aligned} \Pi(z_i, q_i, m_i) &= z_i\mu_i(m_i) + (1 - z_i)\mu_i(1 - m_i) \\ &\quad + \int_0^{z_i} x^A(\theta; q_i)d\theta + \int_0^{r - z_i} x^B(\theta; 1 - q_i)d\theta - q_i w \end{aligned} \tag{3}$$

This expression is the same as in our basic model, except that now the initial ability functions depend on the composition of the lower level, and further the wage differential generates a direct effect on profits from the lower level composition, $-q_i w$.

A key quantity for characterizing the equilibrium is the "mentoring differential" at firm i, which we define as

$$\Delta_i = \mu_i(m_i) - \mu_i(1 - m_i).$$

The mentoring differential of a firm is increasing in the representation of A types in management m_i and in the importance of mentoring.

We now show that type-based mentoring does not necessarily produce segregated firms.

Proposition 10 *(i) Suppose* $m_i > 1/2 > m_j$. *Then the unique equilibrium outcome is* $q_i = 1$, $z_i = r$, $q_j = 0$, $z_j = 0$. *(ii) Suppose* m_1, $m_2 > 1/2$. *Suppose further that:*

$$\frac{\Delta_1^2 + \Delta_2^2}{2|\Delta_1 - \Delta_2|} > r\gamma > \Delta_1. \tag{4}$$

Then there exists an unique interior equilibrium where

$$\begin{aligned} q_i^* &= \frac{1}{2} + r\gamma\left(\frac{\Delta_i - \Delta_{-i}}{\Delta_i^2 + \Delta_{-i}^2}\right) \in (0,1), \\ z_i^* &= q_i^* r + q_1^* q_2^* \Delta_i/\gamma \in (0,r), \\ w^* &= r\left(\frac{\Delta_1\Delta_2^2 + \Delta_2\Delta_1^2}{\Delta_1^2 + \Delta_2^2}\right) > 0. \end{aligned}$$

Part (i) of Proposition 10 states that, if firms start out specialized by type, the unique equilibrium involves full specialization in entry level hiring, and thus full specialization in promotion decisions. This result is due to the fact that if firms are initially specialized, firm 1 values A's more highly than B's while firm 2 values B's more highly than A's. The marginal value of A's to firm 1 (and B's to firm 2) may fall as the lower levels become increasingly segregated, but the ordering of preferences will never reverse. In this case, type-based mentoring leads to segregated firms.

Part (ii) of Proposition 10 shows that if type A is the majority in both firms, then even if one firm is nearly homogenous and the other nearly diverse, it is possible that the firms will both hire and promote some minority workers. This result may seem paradoxical because the mentoring differential varies across the firms while the wage differential is constant. However, the value of hiring a type A depends not just on the mentoring differential, but also on the probability that the lower level employee makes it into management. The more A's a firm hires, the lower the probability that any given A is promoted. Hence, the marginal benefit of hiring A's is diminishing. If the mentoring differentials are not too large, this effect can lead to an interior equilibrium where both firms hire some A's. Equation (4) gives sufficient conditions for the existence of such an interior equilibrium. Essentially, it assures that it is possible to equalize the marginal benefit of hiring A's across firms. Note that as the difference between mentoring differentials, $|\Delta_1 - \Delta_2|$, becomes small, the inequality is always satisfied.[23]

[23]The left hand inequality of (4) has an intuitive interpretation. It guarantees that the marginal value of a type A worker is greater at firm 2 than at firm 1 when all of the type A workers are hired by firm 1 and all of the type B workers are hired by firm 2. The right hand inequality of (4) just assures that the mentoring differential at firm 1 is not so large that $z_1 = r$ and B's are never promoted.

The extent of entry level segregation depends on two factors. First, it is increasing in the difference in mentoring outcomes at the two firms, $|\Delta_1 - \Delta_2|$. Second, segregation is actually *increasing* in the importance of scarcity relative to the magnitude of mentoring differentials, $\gamma/(\Delta_1^2 + \Delta_2^2)$. This somewhat counter-intuitive result arises because the more scarce is initial talent relative to mentoring differentials, the less the probability of promoting a given type A worker falls as more A's are hired, and thus the lower are the opportunity costs of segregation.

As in our study of the basic, model, we can build on these results to consider the implications of labor market interactions for our analysis of the optimal promotion bias of each firm in a two-period model. The labor market introduces asymmetries in the initial ability functions faced by each firm. If $\Delta_1 > \Delta_2$, then the initial ability function of the A's at firm 1 shifts up relative to the B's, while at firm 2 it shifts down (since $q_1^* > 1/2 > q_2^*$). As a result, more A's are promoted at firm 1, and fewer at firm 2, than if the entry level was fully integrated ($q_i = 1/2$). Hence, the bias at firm 1 will tend to shift towards the majority while at firm 2 the bias will shift towards the minority, relative to the single firm case. In the special case where firms are symmetric in terms of their mentoring functions and initial conditions, we will have $q_1^* = q_2^* = 1/2$, and both firms will be integrated at the entry level (though type A workers will receive a wage premium). In this case, our analysis of promotion policies is identical to the single-firm case.

While it is beyond the scope of this paper to solve a fully dynamic, multi-firm model and study whether firms may converge to the single firm outcome, we offer the following result.

Corollary 11 *Suppose $m_1 > m_2 > 1/2$. Then there exists values $\Delta_1 > \Delta_2 > 0$ such that the difference in diversity levels between the two firms decreases relative to the initial conditions as a result of equilibrium hiring.*

Corollary 11 establishes the possibility that despite the pressure for segregation from type-biased mentoring, the equilibrium hiring and promotion decisions will leave firms more similar than their initial conditions. This may happen despite the fact that the more homogenous firm, firm 1, promotes more type A workers ($z_1^* > z_2^*$), since firm 1 also loses more type A workers to retirement.

Observe that our analysis has abstracted from several important considerations. We did not consider that employees may derive a benefit from being promoted. Such a benefit would work in favor of an interior equilibrium.[24] Further, if wage differentials are illegal or create other costs for the firm, the promotion probability might become the driving force for allocating workers to firms.

[24]Suppose A's are in the majority at both firm 1 and 2 and all A's are hired by firm 1. The promotion proability of A's is higher at firm 2. Hence, the greater is the benefit to being promoted, the greater is the wage premium that firm 1 will have to pay to hold on to all of the A's and hence the harder it will be to sustain full segragation as an equilibrium outcome.

6 Ex Ante Human Capital Investment

While the model in the previous section shows that even firms in a competitive labor market may choose a diverse entry level, this result still relies heavily on the availability and scarcity of initial ability among workers of both types. In general, we expect that the relative supply of initial ability will be endogenously determined and thus might not be symmetric.

In particular, we expect types to sort among industries taking into account not only their innate ability in each industry, but also any differentials in wages or promotion probabilities. For example, in the model of the previous section, workers of the majority type receive a wage differential when ability is supplied symmetrically by both types. We might expect additional majority-type workers to enter and minority-type workers to exit (although asymmetric opportunities in other industries may mitigate this effect).

Thus, under some conditions adding a participation or human capital investment decision would increase the relative supply of high-ability majority types in a given industry. This should serve to slow the rate at which firms diversify and lead to lower steady-state levels of diversity. The decisions of firms and workers would reinforce one another: the greater the expected relative quality of the type A worker pool, the more type A workers a firm expects to hire and promote in the future, and the more type A workers the firm wishes to promote today. This would feed back into the initial investment decisions of the workers, encouraging further asymmetries. This might encourage clustering of types in some industries: workers invest in human capital for industries in which their type has at least moderate representation. In fact, such behavior has been observed historically, as is documented in the literature on "sex segregation" (Beller, 1984; Bergmann, 1989). Further, if one type is the minority in most industries, that type might have lower incentives for human capital investment across the board.

Endogenous human capital investment will also have implications for public policy. Our formal analysis involves optimizing behavior without externalities and hence there are no inefficiencies.[25] Unless the social welfare function includes a taste for diversity, our results do not motivate government constraints on the hiring and promotion decisions of firms. However, the strategic interdependence between the decisions of firms and workers can lead to coordination failures. One form of coordination failure is an inefficient level of investment by workers coupled with too little promotion of minorities by firms. In addition, one might see an inefficient concentration of minorities in only a few industries or occupations, when some of those minorities might have scarce

[25]There are other sources of inefficiencies. For example, successful members of some types (i.e. individuals who attain upper level positions) may serve as "role models" for younger members of their community. This role modeling would be very much like our broad definition of mentoring *except that the human capital tranfers would happen outside of the firm* and hence would not be factored into firm decision making.

talent for other industries.[26] Thus, there might be a role for government intervention to increase diversity.

7 Conclusion

This paper develops a novel perspective on discrimination and diversity. We begin with the view that firms possess a stock of upper level employees. With type-based mentoring, the diversity of this stock matters, since it affects the flow of payoffs from newly promoted workers. As a result, a firm's optimal promotion policies have a dynamic component: they affect how the diversity of a firm evolves over time. This leads to a notion of bias as reflecting a firm's willingness to sacrifice current profits to increase or decrease future diversity.

We find that when the firm draws from a diverse applicant pool, its per-period payoff may be maximized anywhere from full diversity to full homogeneity. Thus, our model provides some insight as to why some firms adopt policies of affirmative action, while others oppose such policies. We further show that under some conditions, type-based mentoring may lead to a "glass ceiling," where minority representation reaches a stable steady state which involves less diversity than in the worker pool. Finally, we demonstrate that there may be multiple equilibria. Full diversity may be the optimal stable steady state for a firm, but it may not be optimal for a historically homogenous firm to sacrifice immediate profits to achieve full diversity.

Empirical evidence about the dynamics of diversity and its dependence on current levels of diversity would require data that tracks the diversity of organizations over time. The data on partnership decisions by individual law firms is publicly available, and it would be especially appropriate for investigating the importance of mentoring effects. More broadly, it would be interesting to examine how a variety of individual outcomes (e.g. promotions, hiring, and turnover) depend on the type composition of the firm at different levels of the hierarchy.

In addition, our base model can be extended in a number of theoretical directions. A complete theory of long-run diversity in an economy should endogenize the human capital acquisition decisions of workers, and allow for entry and exit of firms. While we have informally discussed the effects we believe type-based mentoring might have in such a model, a complete theory might formalize these effects as well as further explore the causes and consequences of any market failures and externalities.

On a more micro level, it might be interesting to deepen our understanding of mentoring, and to consider a more general set of assumptions about mentoring. For example, we could study the incentives to build mentoring relationships. In another potential extension, we might allow the level of mentoring received by a worker to

[26] Such concentration could be inefficient if means that minorities are unable to find good matches with their innate talents and interests.

depend on the characteristics of the mentor. If the effectiveness of a mentor depends on her ability, firms would be less willing to bias promotions to achieve a desired level of diversity, since hiring less qualified workers would have future as well as current costs. If the effectiveness of a mentor depends mainly on the mentoring she herself received as an entry-level worker, then initially homogenous firms will face an additional force in favor of inertia. Finally, we might wish to model the game played among workers in the firm's hierarchy, or more generally, the forces which affect the evolution of a firm's corporate culture.

There is also the possibility to further develop a theory of the firm in which firms build up specialized assets over time. This is a view which goes back at least to Prescott and Visscher (1980), but which has not figured prominently in formal theorizing. A variety of decisions which have important effects on a firm's stock of specialized assets can be fruitfully studied from this perspective. For example, firms can be viewed as a stock of competencies whose accumulation depends on the nature of a firm's past activities (such as the set of products it has developed). A natural assumption is that competencies affect a firm's payoff from its current activities (in a process similar to type-based mentoring). Then, a firm's choice of current activities does not just determine its per-period profit, but it has an important dynamic component in that it affects how the firm's competencies evolve over time. The trade-off that arises in this context between breadth and depth of competence appears surprisingly similar to the trade-off between diversity and homogeneity that is the subject of this paper.

8 Appendix

Proof of Lemma 1: In the notation of Stokey and Lucas (1989), our one period payoff function, $F(m,m') = \pi(m, m' - (1-r)m)$ is bounded and continuous; the function $\Gamma(m) = [(1-r)m, (1-r)m + r]$, which characterizes the feasible values of the state variable in the next period, is nonempty, compact valued and continuous; and the discount rate is in $(0,1)$. Hence their Theorem 4.6 holds for our model and the value function is unique and continuous and the policy correspondence is u.h.c. The symmetry in V and $|z^*(m) - r/2|$ follow from the symmetry of our model. □

Proof of Proposition 2:

Recall that $z^{UB}(m;r,\alpha,\gamma) = \arg\max_{z\in[0,r]} \pi(m,z;r,\alpha,\gamma)$, and that $\frac{\partial}{\partial z}\pi = x(z^{UB};\gamma) - x(r-z^{UB};\gamma) + \alpha\left[\widehat{\mu}(m) - \widehat{\mu}(1-m)\right]$. (i) By definition, $x(z^{UB};\gamma) - x(r-z^{UB};\gamma)$ is nonincreasing in γ for $z^{UB} \geq \frac{r}{2}$, which holds if $m \geq \frac{1}{2}$. Thus, $\frac{\partial^2}{\partial z \partial \gamma}\pi \leq 0$, and z^{UB} must be nonincreasing in γ. But then, $(r - z^{UB}(m;r,\alpha,\gamma))/z^{UB}(m;r,\alpha,\gamma)$ is nondecreasing in γ. (ii) Since $\widehat{\mu}(m) - \widehat{\mu}(1-m) > 0$ for $m \geq \frac{1}{2}$, $\frac{\partial^2}{\partial z \partial \alpha}\pi \geq 0$, which implies that $(r - z^{UB})/z^{UB}$ is nonincreasing in α. Thus, SCV is nondecreasing in α. (iii) If $\frac{\partial}{\partial r} z^{UB}(m;r,\alpha,\gamma))\frac{r}{z} \leq 1$, then $(r - z^{UB}(m;r,\alpha,\gamma))/z^{UB}(m;r,\alpha,\gamma)$ is nondecreasing in r. Thus, SCV is nonincreasing in r. □

Proof of Proposition 3:

We proceed by induction on number of periods remaining in a firm's life. We know that when there is 1 period to go, the value function is equal to $\pi^{UB}(m)$, which is nonincreasing in $m \geq \frac{1}{2}$ when SCV holds everywhere. Consider the problem with T periods to go, and assume that the value of the firm with $T-1$ periods to go, $V^{(T-1)}(m)$, is nonincreasing in m. The firm then solves

$$V^{(T)}(m) = \max_z \left\{ \pi(m,z) + \delta V^{(T-1)}((1-r)m + z) \right\}$$

Observe that $\frac{\partial^2}{\partial m \partial z}\pi(m,z) = \mu'(m) + \mu'(1-m) \geq 0$. Since $V^{(T-1)}(m)$ is nonincreasing, clearly $z^{(T)}(m) \leq z^{UB}(m)$, and so $\frac{\partial}{\partial m}\pi(m, z^{(T)}) \leq \frac{\partial}{\partial m}\pi(m, z^{UB}) \leq 0$. Thus, $\frac{d}{dm}V^{(T)}(m) = \frac{\partial}{\partial m}\pi(m, z^{(T)}) + \delta \cdot (1-r)V^{(T-1)'}((1-r)m + z) \leq 0$. By induction, and since monotonicity is preserved by infinite sums, the infinite horizon value function is also nonincreasing in $m \geq \frac{1}{2}$. The argument is analogous for nowhere SCV. □

Proof of Proposition 4:

(i) For $M^{UB}(\cdot)$ nondecreasing, note that $z^{UB}(m) = \arg\max_{z \in [0,r]} \pi(m,z)$ and $\frac{\partial^2}{\partial m \partial z}\pi(m,z) = \mu'(m) + \mu'(1-m) > 0$. Since the objective is differentiable, strictly concave, and strictly supermodular, $z^{UB}(m)$ must be unique everywhere, and strictly increasing when $0 < z^{UB}(m) < r$.

(ii) Consider the following expression for $M^*(m)$

$$M^*(m) = \arg \max_{m' \in \Gamma(m)} \pi(m, m' - (1-r)m) + \delta V(m'),$$

where $\Gamma(m) = [(1-r)m, (1-r)m + r]$. Note that the above objective is strictly supermodular in m and m' since

$$\frac{\partial^2 \pi(m, m' - (1-r)m)}{\partial m \partial m'} = \mu'(m) + \mu'(1-m) - x'(m' - (1-r)m) - x'(r - m' - (1-r)m) > 0. \tag{5}$$

Further, $\Gamma(m)$ is nondecreasing in the strong set order (see Milgrom and Shannon, 1994). Hence, $M^*(m)$ is nondecreasing and single-valued everywhere except for at upward jumps. Thus, every selection from $M^*(m)$ is a continuous function but for upward jumps.

Now, we establish convergence for both M^* and M^{UB}. Clearly $M^{UB}(m > \frac{1}{2}) > \frac{1}{2}$. Consider $m' < \frac{1}{2}$ and $m > \frac{1}{2}$. Condition (5) and the symmetry of V implies that $\pi(m, m' - (1-r)m) + \delta V(m') < \pi(m, 1 - m' - (1-r)m) + \delta V(1-m')$, which establishes that $M^*(m > \frac{1}{2}) > \frac{1}{2}$ for $m > 1/2$. By our constraint, $M^*(m > \frac{1}{2}) \leq 1$. Hence, following Milgrom and Roberts (1994), starting from $m^0 > \frac{1}{2}$, a fixed point of M^* and M^{UB} exists on $[\frac{1}{2}, 1]$. □

Proof of Proposition 5:

(i) $\frac{1}{2} \in S^{UB}$ if and only if $z^{UB'}(\frac{1}{2}) \leq r$. Using the implicit function theorem, $z^{UB'}(\frac{1}{2})$ $=-\ \mu'(\frac{1}{2})/x'(\frac{r}{2})$; substituting gives the result.

(ii) Consider $x(\theta) = 1 - \theta$ and $\mu(m) = rm$. Then $z^{UB}(m) = rm\ \forall m$ and $S^{UB} = [0, 1]$. It is then straightforward to perturb this x or μ so that $z^{UB}(m) \neq rm$ for $m \in N$, where N is any subset of $(\frac{1}{2}, 1]$ which contains no isolated points. □

Example 1: Suppose $x(\theta) = \kappa - \gamma\theta$. Since $M^{UB}(m) = (1 - r)m + z^{UB}(m)$, the curvature of M^{UB} follows that of z^{UB}. For $z^{UB} < 1$,

$$z^{UB}(m) = \frac{r}{2} + \frac{\mu(m) - \mu(1-m)}{2\gamma}$$

$$\Rightarrow \frac{\partial^2 z^{UB}}{\partial m^2} = \frac{\mu''(m) - \mu''(1-m)}{2\gamma}$$

Hence, if $\mu'''(m) > 0$ for all m, then M^{UB} is concave over $(\frac{1}{2}, 1]$ and there is at most one stable steady state in this interval; the steady state exists iff $M'(\frac{1}{2}) > 1$. If $\mu'''(m) < 0$ for all m, then M^{UB} is convex over $(\frac{1}{2}, 1]$ and there are at most two steady states in the interval, and only state $m = 1$ can be stable. □

Proof of Proposition 6: (i) Follows by Milgrom and Roberts (1994), using our results on the sign of the bias from Proposition 3. (ii) $z^{UB}(\frac{1}{2}) = \frac{r}{2}$. Hence, $z^*(\frac{1}{2}) = z^{UB}(\frac{1}{2}) + b(\frac{1}{2}) = \frac{r}{2}$ iff $b(\frac{1}{2}) = 0$, which is the case when SCV holds everywhere, but not when SCV nowhere.

Proof of Lemma 8: Recall that Lemma 4 implies that M^* is single-valued almost everywhere, except at upward jumps. By definition, rm_s is one optimal promotion policy at state m_s. Suppose that there is another optimal promotion rule, $z' > rm_s$. Then there must be an upward jump in M^* at m_s; suppose that the jump is of size d. But this implies that for all $d > \varepsilon > 0$, $M^*(m_s + \varepsilon) > m_s + d > m_s + \varepsilon$, contradicting the definition of a stable steady state.

Now consider an interior steady state $m_s \notin \{0, \frac{1}{2}, 1\}$ and the associated steady state promotion policy $z_s = rm_s$. Since

$$z_s \in \arg\max_{z \in [0,r]} \pi(m_s, z) + \delta V(z + (1 - r)m_s)$$

it must be that $\pi_2(m_s, rm_s) + \delta V'(m_s) = 0$ whereever this derivative exists. To solve for V', we write out the value function in long form as the sum of per period payoffs for the vector of promotion decisions $\vec{z} = (z^1, z^2, \ldots)$. That is, define

$$\tilde{V}(m, \vec{z}) = \pi(m, z^1) + \delta\pi(z^1 + (1 - r)m, z^2) + \ldots$$

If m_s is a steady state, then $V'(m_s) = \frac{d}{dm_s}\tilde{V}(m_s, \vec{z}_s)$ where $\vec{z}_s = (rm_s, rm_s, \ldots)$. By the envelope theorem, we have that

$$\frac{d}{dm_s}\tilde{V}(m_s, \vec{z}_s) \quad = \quad \frac{\partial}{\partial m}\tilde{V}(m_s, \vec{z}_s)$$

$$= \sum_{t=0}^{\infty}(1-r)^t\delta^t\pi_1(m_s, rm_s)$$
$$= \frac{\pi_1(m_s, rm_s)}{1-(1-r)\delta}.$$

Inserting this into the expression $\pi_2(m_s, rm_s) + \delta V'(m_s) = 0$ gives the result.

Proof of Proposition 9: (ii) Following the arguments from the text, we apply the implicit function theorem. Using

$$\pi_1(m_s, rm_s) = x(rm_s) - x(r(1-m_s)) + \mu(m_s) - \mu(1-m_s)$$
$$\pi_2(m_s, rm_s) = m_s\mu'(m_s) - (1-m_s)\mu'(1-m_s)$$

we have that

$$\frac{\partial FOC_s(m;\delta)}{\partial\delta} = \frac{r(m\mu'(m) - (1-m)\mu'(1-m))}{(1-(1-r)\delta)^2}.$$

Consider $m > \frac{1}{2}$. If μ is globally SCV, then $z^{UB}(m)\mu'(1-m) > (r - z^{UB}(m))\mu'(m)$. Then $z^{UB}(m) > z^*(m) = rm$ implies that $\partial FOC_s(m)/\partial\delta > 0$. Conversely, if μ is nowhere SCV, $\partial FOC_s(m)/\partial\delta < 0$. □

Proof of Proposition 10: Suppose that m_1, $m_2 > 1/2$ and that *w.l.o.g* $\Delta_1 \geq \Delta_2$. In an interior equilibrium

$$\frac{\partial\Pi}{\partial z_i} = \mu_i(m_i) - \gamma z_i/q_i - \mu_i(1-m_i) + \gamma(r-z_i)/(1-q_i) = 0$$

$$\Rightarrow z_i = q_i(1-q_i)\Delta_i/\gamma + q_i r$$

We can use this expression for z_i to write the first order condition for q_i as

$$\frac{\partial\Pi}{\partial q_i} = \frac{\gamma z_i^2}{2q_i^2} - \frac{\gamma(r-z_i)^2}{2(1-q_i)^2} - w = 0,$$

$$\Rightarrow q_i(w) = 1/2 + r\gamma/\Delta_i - w\gamma/\Delta_i^2.$$

Imposing the equilibrium condition $q_1(w^*) + q_2(w^*) = 1$, yields the expression for w^*, which can then be used to solve for q_i^* and z_i^*. These quantities constitute an interior equilibrium if the following two conditions hold:

$$1 > q_1^* \Leftrightarrow \frac{\Delta_1^2 + \Delta_2^2}{2(\Delta_1 - \Delta_2)} > r\gamma,$$

$$r > z_1^* \Leftrightarrow r\gamma > \Delta_1.$$

Suppose that $m_i > 1/2 > m_j$. There cannot be an interior equilibrium since $w \geq 0 \Rightarrow q_j = 0$ and $w \leq 0 \Rightarrow q_i = 1$. Hence the unique equilibrium outcome must

be $q_i = 1$, $z_i = r$, $z_j = 0$. Price must be set sufficiently close to zero that the market clears. □

Proof of Corollary 11: Consider two values of Δ_1 and Δ_2 which satisfy (4) and $\Delta_1 > \Delta_2 > 0$. The diversity of firm i at the end of the period is $(1-r)m_i + z_i^*$. The firms become more similar in their diversity levels if

$$(1-r)m_1 + z_1^* - ((1-r)m_2 + z_2^*) < m_1 - m_2,$$

which is equivalent to

$$(\Delta_1 - \Delta_2)\left(\frac{1}{4r\gamma} + r\gamma\left(\frac{\Delta_1 + \Delta_2}{\Delta_1^2 + \Delta_2^2}\right)^2\right) < m_1 - m_2.$$

As Δ_2 approaches Δ_1, the latter inequality will be satisfied, and (4) will continue to hold. Thus, if the firms are similar enough at the start, they will become more similar.□

9 References

Arrow, Kenneth (1973). "The Theory of Discrimination," in O. Ashenfelter and A. Rees (eds) *Discrimination in Labor Markets*, Princeton University Press.

Barlett, Robin, 1997, "Report of the Committee on the Status of Women in the Economics Profession," *American Economic Review* 87(2): 506-510.

Bergmann, Barbara (1989). "Does the Market for Women's Labor Need Fixing?" *Journal of Economic Perspectives*, 3, Winter, 43-60.

Beller, Andrea (1984). "Trends in Occupational Segregation by Sex and Race, 1960-1981," in *Sex Segregation in the Workplace*, edited by Barbara F. Reskin, Washington, DC: National Academy.

Carmichael, Lorne (1988). "Incentives in Academics: Why Is There Tenure?" *Journal of Political Economy* 96 (3): 453-472.

Coate, Stephen and Loury, Glenn (1993). "Will Affirmative-Action Policies Eliminate Negative Stereotypes?" *American Economic Review* December, 1220–1240.

Cain, Glen (1986). "The Economic Analysis of Labor Market Discrimination: A Survey," in *The Handbook of Labor Economics*, edited by Orley Ashenfelter and Richard Layard, New York: Elsevier, Science.

Cornell, Bradford and Ivo Welch (1996). "Culture, Information and Screening Discrimination," *Journal of Political Economy*104 (3):542-572.

Dreher, George F., and Ash, Ronald A. (1990), "A comparative study of mentoring among men and women in| managerial, professional, and technical positions." *Journal of Applied Psychology* 75 (October), 539-46.

Dreher, George F., and Cox, Taylor H. (1996), "Race, gender, and opportunity: a study of compensation attainment and the establishment of mentoring relationships." *Journal of Applied Psychology* 81 (June) pp. 297-308.

Gilson, Ronald and Mnookin, Robert (1989). "Coming of Age in a Corporate law Firm: The Economics of Associate Career Patterns," *Stanford Law Review* 41 (February): 567-595.

Greif, Avner (1994), "Cultural Beliefs and the Organization of Society: A Historical and Theoretical Reflection on Collectivist and Individualist Societies." *Journal of Political Economy* 102 (5) October, pp. 912-950.

Greif, Avner (1993), "Contract Enforceability and Economic Institutions in Early Trade: the Maghribi Traders' Coalition," *American Economic Review* 83 (3) June, pp. 525-48.

Ibarra, Herminia (1992), "Homophily and differential returns: sex differences in network structure and access in an advertising firm." *Administrative Science Quarterly* 37 (September) pp. 422-47.

Kanter, Rosabeth (1977). *Men and Women of the Corporation*, New York: Basic Books, Inc.

Laband, David N., and Lentz, Bernard F. (1994), "Workplace mentoring in the legal profession," *Southern Economic Journal* 61 (January) pp. 783-802.

Milgrom, Paul and Roberts, D. John (1994). "Comparing Equilibria," *American Economic Review*, June, 441–459.

Milgrom, Paul and Shannon, Christina (1994). "Monotone Comparative Statics," *Econometrica*, 60 (1), p. 157-80.

Morrison, Ann, and Von Glinow, Mary Ann (1990). "Women and Minorities in Management." *American Psychologist*, February, 200–208.

Noe, Raymond (1988). "Women and Mentoring: A Review and Research Agenda." *Academy of Management Review* , 13 (1): 65-78.

Orpen, Christopher (1995). "The effects of mentoring on employees' career success." *The Journal of Social Psychology* 135 (October) pp.. 667-8.

Prescott Edward and M. Visscher (1980). "Organizational Capital." *The Journal of Political Economy* 88: pp. 446–461.

Riley, Sandra and Wrench, David (1985), "Mentoring among women lawyers." *Journal of Applied Social Psychology* 15 (4): pp. 374- 86.

Rosen, Asa (1997). "An Equilibrium Search-Matching Model of Discrimination." *European Economic Review*, 41: 1589–1613.

Rothstein, Donna (1997). "Supervisor Gender, Race, and Ethnicity and the Labor Market Outcomes of Young Workers." Mimeo, U.S. Bureau of Labor Statistics.

Spurr, Stephen, 1990, "Sex Discrimination in the Legal Profession: A Study of Promotion." *Industrial and Labor Relations Review* 43 (April): 406-417.

Stokey, Nancy, and Lucas, Robert (1989). *Recursive Methods in Economic Dynamics*, Cambridge: Harvard University Press.

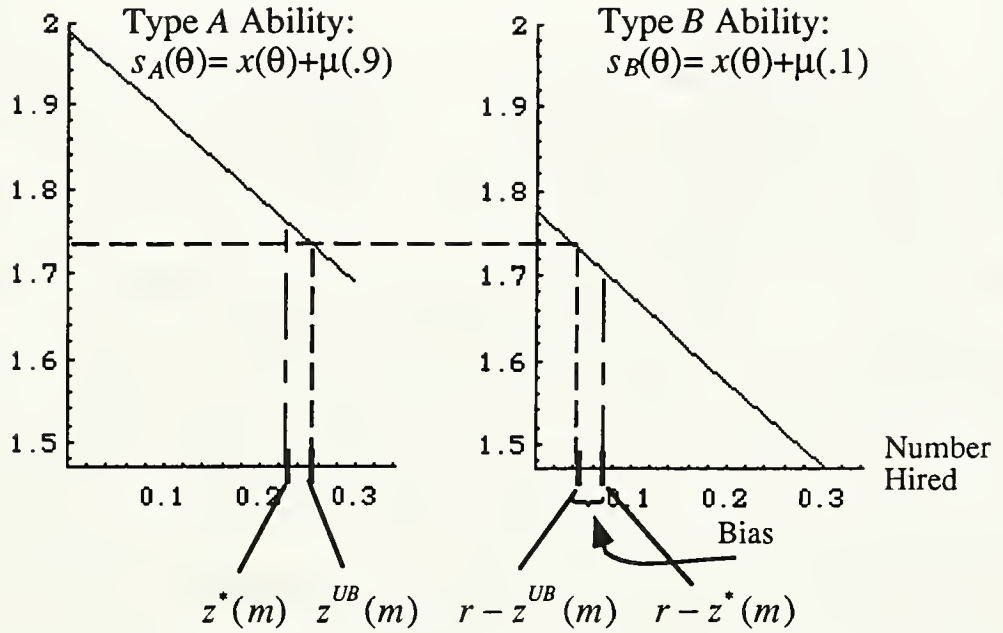

Figure 1: The figure plots surplus functions and promotion policies. The ability of type *A* workers is higher than type *B* due to the mentoring differential. As indicated with the horizontal line, unbiased promotions equalize final abilities across the two types. The optimal promotion policy for this set of parameters entails a bias in favor of the minority, type *B*.
Parameters: $\mu(m) = m^{.11}$, r=.3, $x(\theta) = 1 - \theta$, m=.9, δ=.95.

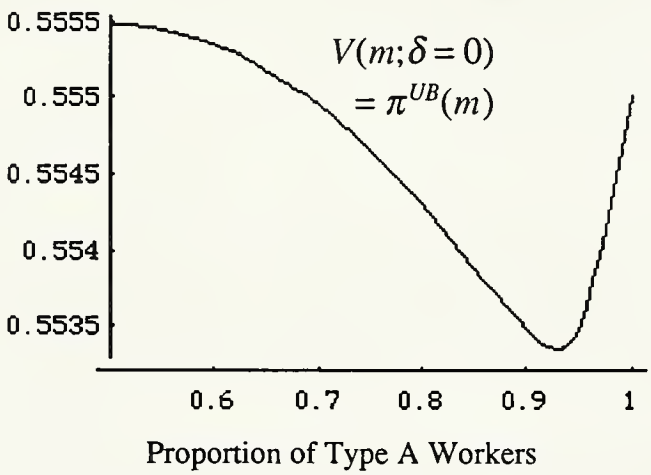

Figure 2: One-period value function, δ=0.
Parameters: $\mu(m) = m^{.11}$, r=.3, $x(\theta) = 1 - \theta$.

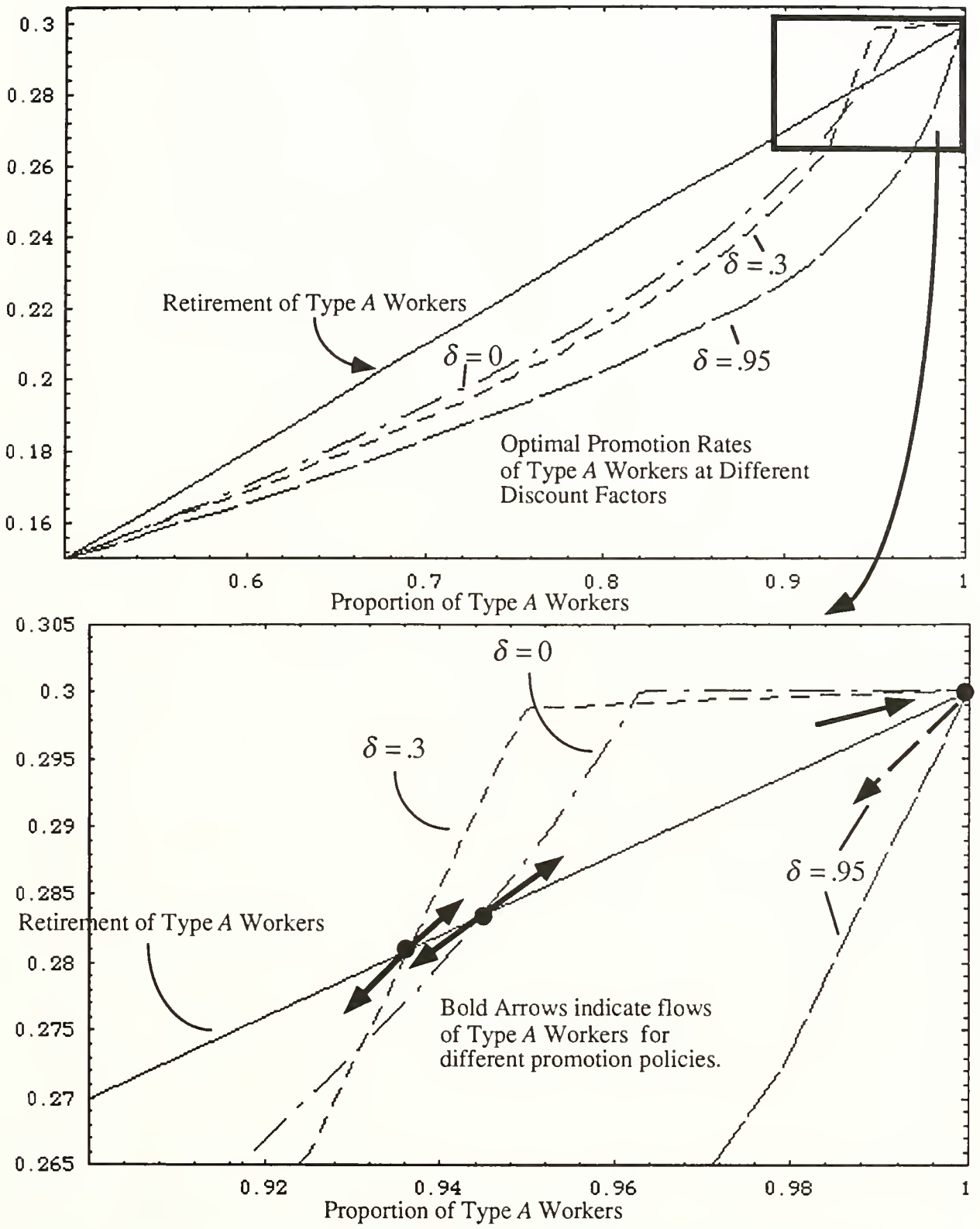

Figure 3: The dynamic evolution of diversity for different discount factors. The upper panel plots the retirement rate against the optimal promotion policies for δ=0, .3, and .95. As indicated by the arrows in the lower panel, when the retirement rate is greater (less) than the promotion rate, the firm moves towards greater (less) diversity. Parameters: $\mu(m) = m^{.11}$, r=.3, $x(\theta)$=1−θ.

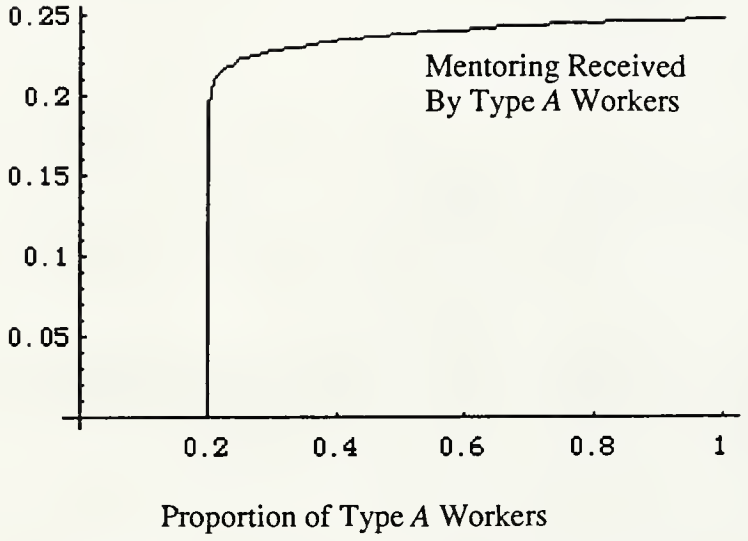

Figure 4: A critical mass mentoring function, where $\mu(m) = 0$ when $m<.2$, $\mu(m) = .15 + .1(\mu-.2)^{.11}$ otherwise.

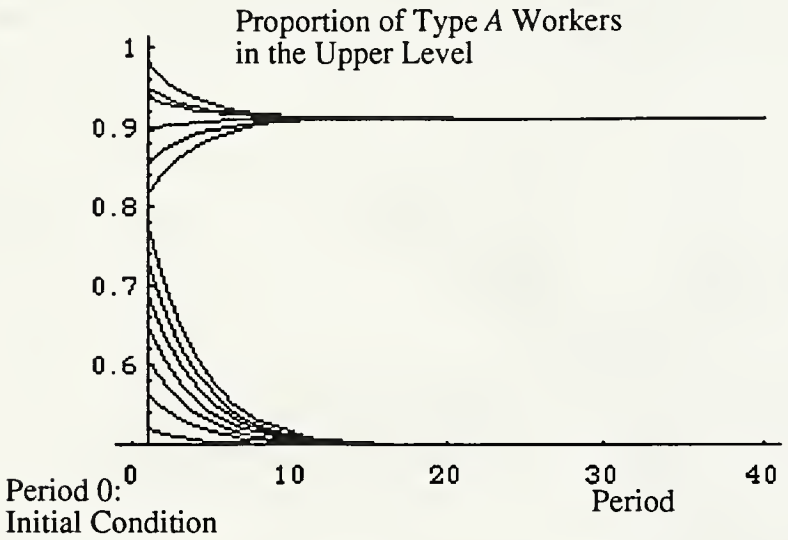

Figure 5: An illustration of the "Glass Ceiling." The figure plots the state of the firm as a function of time, starting from different initial conditions. Mentoring function is critical mass mentoring function from Figure 4, and optimal promotions are used. Parameters: r=.3, δ=.3, $x(\theta)$=1−θ.

7076 .01

Date Due

Lib-26-67

MIT LIBRARIES
3 9080 01972 0728